REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
29-Sep-2006

2. REPORT TYPE
REPRINT

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
ADVANCES IN MIXED SIGNAL PROCESSING FOR REGIONAL AND TELESEISMIC ARRAYS

5a. CONTRACT NUMBER
FA8718-04-C-0013

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62601F

6. AUTHOR(S)
Robert H. Shumway

5d. PROJECT NUMBER
1010

5e. TASK NUMBER
SM

5f. WORK UNIT NUMBER
A1

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Statistics
University of California at Davis
1 Shields Avenue
Davis, CA 95616

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/VSBYE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-VS-HA-TR-2006- 1161

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES
Reprinted from: Proceedings of the 28th Seismic Research Review – Ground-Based Nuclear Explosion Monitoring Technologies, 19 – 21 September 2006, Orlando, FL, Volume I pp 503 - 509.

14. ABSTRACT

In this project, we have considered a number of possible approaches to resolving mixtures of propagating signals and noises over seismic arrays. The problem is important because there will often be interfering noise sources, extraneous signals, or mixtures of phases that will be organized as plane waves. Analyzing such mixtures using conventional beam-forming and other methods will produce distorted velocities and azimuths and may even give incorrect magnitudes. One interesting result of this study is that beam-forming, single-signal F-Statistics as well as typical high resolution multiple-signal estimators from engineering literature, such as the Multiple Signal Characteristic (MUSIC) algorithm all suffer from the above problems for some very common situations involving regional and teleseismic data.

We have approached the problem by showing that the multiple-signal plane-wave model is essentially in the form of a nonlinear regression in the frequency domain so that a sequential approach to isolating component signals analogous to stepwise linear regression can be applied. This approach provides a sequence of tests showing the power contribution of each component of a mixture and yields estimators for velocities and azimuths and their uncertainties. In the first year, we used this regression model, coupled with a corrected form of Akaike's Information Criterion (AICC) to settle on a final mixture. Applying the AICC as a model selection criterion yielded three signals known to have occurred on this long-period mixture and a decomposition of a regional earthquake into a known signal and a propagating noise.

In the second year of the contract, we have concentrated on deconvolving the mixtures obtained using the inverse Fourier components of the signals isolated in the above nonlinear regression model. Applying the approach to the long period mixture previously analyzed produces estimated waveforms for the two known simultaneously occurring earthquakes and for a noise component originating from South Pacific storms. We also show an analysis of a regional event from China that contains a signal and a strong propagating noise component. Deconvolving the mixture produces a waveform for the signal that shows a potential depth phase that cannot be seen in the original waveforms.

15. SUBJECT TERMS
Seismic characterization, Seismic propagation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert J. Raistrick
UNCLAS	UNCLAS	UNCLAS	SAR	7	19b. TELEPHONE NUMBER *(include area code)* 781-377-3726

ADVANCES IN MIXED SIGNAL PROCESSING FOR REGIONAL AND TELESEISMIC ARRAYS

Robert H. Shumway

University of California, Davis

Sponsored by Air Force Research Laboratory

Contract No. FA8718-04-C-0013

ABSTRACT

In this project, we have considered a number of possible approaches to resolving mixtures of propagating signals and noises over seismic arrays. The problem is important because there will often be interfering noise sources, extraneous signals, or mixtures of phases that will be organized as plane waves. Analyzing such mixtures using conventional beam-forming and other methods will produce distorted velocities and azimuths and may even give incorrect magnitudes. One interesting result of this study is that beam-forming, single-signal F-Statistics as well as typical high resolution multiple-signal estimators from engineering literature, such as the Multiple Signal Characteristic (MUSIC) algorithm all suffer from the above problems for some very common situations involving regional and teleseismic data.

We have approached the problem by showing that the multiple-signal plane-wave model is essentially in the form of a nonlinear regression in the frequency domain so that a sequential approach to isolating component signals analogous to stepwise linear regression can be applied. This approach provides a sequence of tests showing the power contribution of each component of a mixture and yields estimators for velocities and azimuths and their uncertainties. In the first year, we used this regression model, coupled with a corrected form of Akaike's Information Criterion (AICC) to settle on a final mixture. Applying the AICC as a model selection criterion yielded three signals known to have occurred on this long-period mixture and a decomposition of a regional earthquake into a known signal and a propagating noise.

In the second year of the contract, we have concentrated on deconvolving the mixtures obtained using the inverse Fourier components of the signals isolated in the above nonlinear regression model. Applying the approach to the long period mixture previously analyzed produces estimated waveforms for the two known simultaneously occurring earthquakes and for a noise component originating from South Pacific storms. We also show an analysis of a regional event from China that contains a signal and a strong propagating noise component. Deconvolving the mixture produces a waveform for the signal that shows a potential depth phase that cannot be seen in the original waveforms.

OBJECTIVES

This project is aimed at applying recently modified array signal processing techniques to problems involving single and multiple signals observed on teleseismic and regional arrays. We are focused on Topic 3 (Seismic Detection, Location and Discrimination) with particular emphasis on proposing "new techniques to enhance signal detection and parameter estimation (e.g., azimuth, phase velocity) in strongly heterogeneous media."

Specifically, we have developed the sequential F-statistic and model selection criterion AICC as off-line techniques for determining the number of signals in a possible mixture of propagating signals and noises. This includes specifying estimators for velocities and azimuths of the components along with their estimated uncertainties. Finally, the estimated velocities and azimuths enable deconvolutions producing separate estimators for each of the component waveforms. A set of MATLAB subroutines has been developed that can be merged into off-line software for investigative research purposes.

RESEARCH ACCOMPLISHED

Several algorithms, such as the sequential F-detector considered here and the MUSIC algorithm, are available that offer promise for handling array data with low signal-to-nose ratios and contamination from interfering signals. In this project, we have first investigated the performance of currently available algorithms on teleseismic and regional data containing mixed signals in order to demonstrate the superior performance of the sequential F-statistic. A sequential analysis of power using the F-statistic is employed that estimates the correct number of signals and their velocities and azimuths. This is contrasted with results using conventional F-K estimators that do not handle the mixed signal case.

Approaches to detecting signals on arrays all focus on the basic model that expresses the observed channel as sums of delayed signals and a unique noise process. The delays are functionally dependent on velocity and azimuth if the signals are propagating plane waves (this is the assumption that is usually made). Methods that are commonly in use for analyzing such data when a single signal is assumed to be present can be roughly categorized as (1) beam-forming and plotting the power as a function of slowness, which can be converted to estimators of velocity and azimuth, (2) beam-forming converted to an F-statistic by dividing by an estimator of the noise power (see Shumway, 1983; Shumway et al., 1999; Blandford, 2002, a, b) (3) Capon's estimator (see Capon, 1969), (4) MUSIC (Schmidt, 1979; Stoica and Nehorai, 1989) and (5) cross correlation (Tribuleac and Herrin, 1997). Only the MUSIC estimator listed above seems to be at all appropriate for analyzing the mixed signal case. The first four of the above estimators are available as the MATLAB subroutine *beam_sl.m*.

In the first year of this contract, we have concentrated on developing the multiple signal F-detector as an improvement over conventional detectors, such as those given above or the currently favored ratio of short term to long term mean squares (STA/LTA). The multiple signal F-detector fits successively higher order nonlinear regression models as the potential number of propagating signals or noises are increased. The number of propagating signals present is selected using the corrected information theoretic criterion of Hurvich and Tsai (1989). The software is available as the MATLAB subroutine *mul_sig_sl.m*. Technical difficulties in applying the Cramer lower bound approach to getting the variances has necessitated the use of the frequency domain bootstrap (Paparoditis and Politis, 1999) as an alternative. Software in the form of the MATLAB subroutine *mul_boot_sl.m* now computes the detection results along with standard deviations and confidence intervals for velocity and azimuth from the bootstrap distribution for data containing an arbitrary number of signals. Empirical results indicate that the distribution is approximately normal but this is not necessary for the validity of the bootstrap confidence intervals.

Analysis with Conventional Detectors

Conventional methods, such as 1–5 above, have met with varying degrees of success when they have been applied in practice in cases where there are known to be interfering signals. We can illustrate some of the pitfalls by considering the two events shown in Figures 1 and 2. Both events in this section were analyzed using the MATLAB software *beam_sl*.

Figure 1 shows one channel from the seven channels containing a mixture of two simultaneously occurring earthquakes, one from South Africa and the other from the Philippines, observed at the USAEDS long-period seismic array in Korea. The correct back-azimuths for these events are 226 degrees and 198 degrees. Yet simple

time delay estimation for this event gives a back-azimuth of 203 degrees, which is closer to the second of the above two signals but still off by 5 degrees. Beam forming or the F-statistic (equivalent to semblance) also produces an estimated azimuth of 203 degrees with a 95% confidence interval (200, 206), not including either of the known signals. The Capon and MUSIC estimators gave 207 and 212 degrees respectively, still between the two correct azimuths. However the analysis using AICC and the sequential F-detector in the next section shows that there are actually three signals (see Tables 1 and 2) in this mixture at 223 (South Africa) degrees, 200 (Philippines) degrees and 130 degrees (South Pacific storm).

Figure 2 shows a noisy event from China, namely the well-located event 1991101_1325, used by Stroujkova and Reiter (2006, these Proceedings). In this case, estimated azimuths were 284, 283, and 288 degrees respectively with velocities 8.1, 10.8 and 10.5 km/s, respectively. All detectors indicating a single signal used the full 50 s of data to retain sufficient degrees of freedom over the relatively narrow bandwidths for the signal (2Hz). The Capon and MUSIC estimators indicated second signals at 291 and 286 degrees with lower velocities (6.2, 6.6 km/s), probably corresponding to a later phase. An analysis of the noise preceding the signal over a lower frequency range indicated two noises, one at about 130 degrees and one at about 236 degrees using all methods. We will find in the next section that the signal band includes two signals (see Tables 3 and 4), one at 286 degrees and one at 148 degrees with relatively high confidence.

Analysis with the Multiple Signal F-Detector

The analysis of multiple signals involves considering a succession of nonlinear regression models written in the frequency domain with parameters expressed in terms of slowness. Beginning with the single-signal model with an estimated set of slowness coordinates, we consider an alternative model with two signals. The likelihood ratio test for a two-signal model against a single-signal model yields a monotone function of an F-statistic. The numerator is the difference between the error power under the two-signal model and the error power under the single-signal model and represents the reduction in power possible from the added signal. This reduction is scaled by the noise power under the full model and a function of the number of parameters and the error degrees of freedom. The sequential fitting of more signals continues one at a time until no more added signals are statistically significant. The final estimated velocities and azimuths are those obtained under the best model. Because the F-statistic at each stage involves the nonlinear parameter slowness, we use the corrected AIC, say AICC of Hurvich and Tsai (1989). The computations in this section were done using the MATLAB subroutine ***mul_sig_sl.m***.

Standard errors and confidence intervals for velocities and azimuths are computed using the frequency domain bootstrap of Paparoditis and Politis (1999) adapted to the nonlinear regression case under the multiple-signal model. This involves reconstructing the frequency domain observations from the regression model evaluated at the maximum likelihood estimators for slowness. The residuals from this model will be roughly independent and constitute the basic resampling population. To reconstruct a bootstrap sample of the data, draw a sample of these residuals with replacement and use the nonlinear regression model to reconstitute a pseudo-sample of the observed data. The estimated velocity and azimuth computed from this pseudo-sample constitute the first pair of estimated parameters. Repeat the above procedure a large number of times (500) and retain the estimators. The sampling distribution of these estimators yields the standard deviations and the 95% confidence intervals shown in Tables 1–4 below. The bootstrap software can be accessed through the subroutine ***mul_boot.m***.

Tables 1 and 2 show the results of the sequential F-tests applied to the long period event and give the confidence intervals resulting from the best model. Note that the first signal identified at azimuth 203 degrees accounts for 85% of the total power and still gives an estimated azimuth midway between the two known azimuths of the mixture. The F-detector is highly significant. Adding a second signal to the model substantially increases the percentage of power accounted for and still yields a highly significant F. Testing for the third signal again produces a highly significant F-statistic and increases the power accounted for to 98%. The fourth potential signal adds a negligible amount to the power accounted for and shows a fairly small F value which is still significant. However, because it only accounts for 1% additional power and the AICC values are minimized for the three-signal model, we take that model as our final configuration.

Table 1. Analysis of power for long period mixture. Sequential F-tests and AICC model selection.

Source	Added Power	F-Statistic	AICC	% Power
First Signal	396	36.3	-1.23	86
Second Signal	43	9.9	-1.86	95
Third Signal	14	7.9	-2.23	98
Fourth Signal	3	2.2	-1.59	99

Table 2 shows the estimated azimuths for the three component signals as 200, 223, and 130 degrees. The first two match up well with the known values but the third has not been identified from alternate records as a real seismic signal. Possibly this third signal is a coherent noise source; there was a known storm in the South Pacific at this time which corresponds to the third azimuth. Confidence intervals include the true azimuths for the two known earthquakes.

Table 2. Estimated velocities and azimuths for single-signal and best model for long period mixture. True azimuths are 226 and 198 degrees respectively. Confidence intervals are from the bootstrap distribution.

Model	Azimuth(Velocity)	95% Azimuth	95% Velocity
Single-Signal	203(3.85)	(200, 206)	(3.8, 3.9)
Three-Signal	200(3.5)	(196, 203)	(3.4, 3.6)
	223(3.9)	(218, 228)	(3.8, 4.0)
	130(4.0)	(125, 133)	(3.8, 4.1)

Tables 3 and 4 repeat the analysis on the regional earthquake from China recorded at the Korean Seismic Array (KSAR) analyzed by Stroujkova and Reiter (2006, these Proceedings) who noted the substantial noise present in the P and pP windows at approximately 150 degrees. The first channel for the China event, shown in Figure 2, exhibits the typical nature of regional recordings characterized by high noise levels and multiple arrivals. In this case, substantially less power is accounted for by the signal components than in the teleseismic case. The F-statistic and AICC settle on the two-signal model with the F-statistic marginally significant for the third signal. The third signal adds little power (4%) to the power accounted for and there is a clear minimum in the model selection criterion AICC at the two-signal model. The total power accounted for by the signal and noise only amounts to 54%.

Table 4 gives the estimated azimuths and velocities for the China event, and we note agreement between the single-signal and two-signal models. Note the smaller velocity for the noise component, namely in the neighborhood of 3 km/s. This is less than half that of the primary signal and may indicate that the noise is propagating more slowly than the primary signal. Note that this result agrees with the noise analysis preceding the signal that indicated estimated velocities in the 3–5 km/s range.

Table 3. Analysis of power for regional mixture. Sequential F-Tests and AICC model selection.

Source	Added Power	F-Statistic	AICC	% Power
First Signal	251	7.5	-.0485	32
Second Signal	138	5.2	-.2067	50
Third Signal	.30	1.2	-.1223	54

Table 4. Estimated velocities and azimuths with 95% confidence intervals for China event.

Model	Azimuth(Velocity)	95% Azimuth	95% Velocity
Single-Signal	286.1(8.1)	(284.9, 287.9)	(8.0, 8.3)
Two-Signal	286.4(8.1)	(285.2, 287.9)	(7.9, 8.2)
	148.1(3.0)	(145.9, 149.7)	(2.9, 3.2)

Deconvolution of Component Signals and Noises

Figures 1 and 2 give the estimated signals that follow from the velocities and azimuths implied by Figures 1–4. The deconvolutions are done using the MATLAB subroutine *mul_decon.m*.

The most prominent signal is from South Africa (228 degrees) and clearly shows in the region between 800 and 11,400 points (30 s). The smaller signal from the Philippines (196 degrees) is less prominent and appears to model parts not accounted for by the first signal. The noise component fills in the region before the first signal and accounts for the burst at the end of the data.

The second deconvolution corresponding to the two components noted in the previous section gives the estimated signal and noise components in the regional data. The estimated noise component from 148 degrees shows the character of the low frequency noise before the signal enters. The estimated signal from 286 degrees shows the depth phase pP much more clearly than does the single channel mixture. In this case, one obtains from the plot a delay of about 4.6 s, which is comparable to that obtained by Stroujkova and Reiter (2006). We also note the noise reduction capabilities of the two-signal beam.

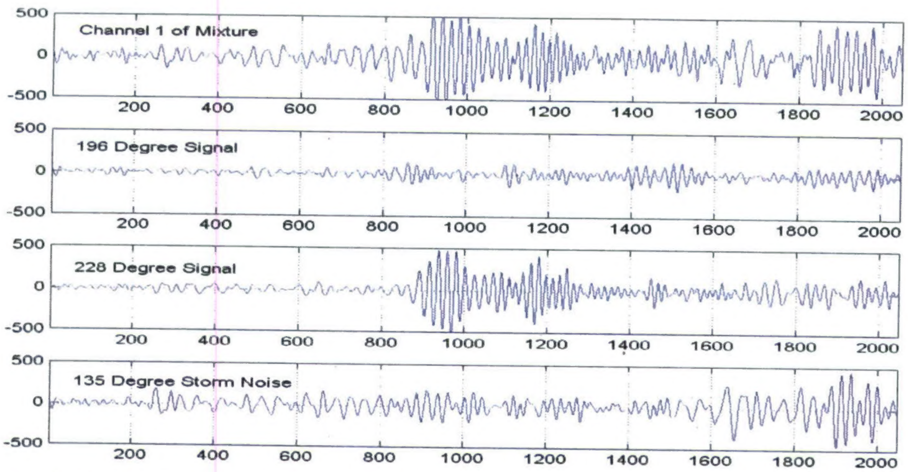

Figure 1. Deconvolution for long-period mixture of two signals and a noise source.

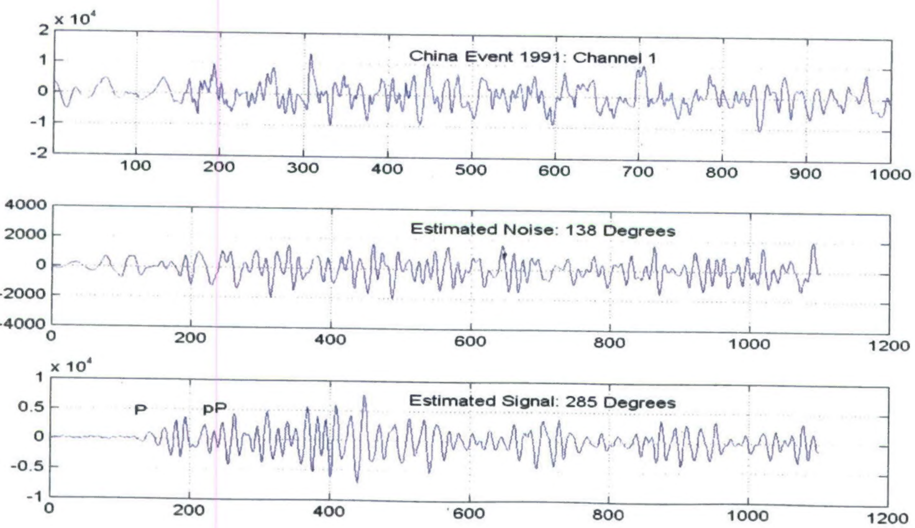

Figure 2. Deconvolution of 1991 China event containing a signal from 285 degrees and an apparent noise source.

CONCLUSIONS AND RECOMMENDATIONS

This work was primarily motivated by the problem of detecting mixtures of signals on teleseismic and regional arrays. Such undetected mixtures are shown to give incorrect velocities and azimuths when treated by traditional single-signal detection methods based on cross-correlation or frequency wave-number methods. Furthermore, the separation of interfering phases and coherent noise sources should improve detection statistics and lead to improvements in location and magnitude estimates.

Using current improved computing platforms, such as MATLAB, makes the nonlinear estimation problems implicit in multiple-signal modeling tractable and easy to implement. Using early work involving multiple-signal estimation (Shumway, 1970) and its extensions to wave-number methods (Smart, 1972, 1976), we are able to formulate the problem in a regression framework that leads to a sequential detection approach for reliably determining the number of signals and coherent noises present along with their estimated velocities and azimuths. We have provided methods for comparing detection performance using the F-statistic, the AICC model selection statistic, and confidence intervals for the velocities and azimuths using the bootstrap. Finally, we have show that the regression model provides a method for deconvolving the component signals and have shown results for both long and short period seismic data.

ACKNOWLEDGEMENTS

Gene Smart and Dean Clauter of the Air Force Technical Applications Center (AFTAC) have graciously provided the data used in this analysis as well as valuable insights into potential methods of analysis. In particular, the previous seminal work of Smart (1972, 1976) is gratefully acknowledged. I am grateful to Anastasia Stroujkova and Delaine Reiter of Weston for providing the data from the regional Chinese event and making available their preprint.

REFERENCES

Blandford, R. R. (2002a). Detection and azimuth estimation by infrasonic arrays as a function of array aperture and signal coherence, Air Force Technical Applications Center AFTAC-TR-02-007.

Blandford, R. R. (2002b). A plan of development for detection systems for seismic and infrasound arrays. Air Force Technical Applications Center AFTAC-TR-02-005.

Capon, J. (1969). High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE*, 57: 1.408–1.418.

Hurvich, C. M. and C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika* 76: 297–307.

Paparoditis, E. and D. N. Politis (1999). The local bootstrap for periodogram statistics. *J. Time Series Anal.* 20: 193–222.

Schmidt, R. O (1979). Multiple emitter location and signal parameter estimation, in *Proceedings of the RADC Spectral Estimation Workshop*, pp. 243–258.

Shumway, R. H. (1970). Applied regression and analysis of variance for stationary time series. *J. Amer. Statist. Assoc.* 65: 1,527–1,546

Shumway, R. H. (1983). Replicated time series regression: An approach to signal estimation and detection. *Handbook of Statistics 3: Time Series in the Frequency Domain*, D. R. Brillinger and P. R. Krishnaiah (Eds.). Ch. 18, pp. 383–408.

Shumway, R. H., S. E. Kim and R. R. Blandford (1999). Nonlinear estimation for time series observed on arrays. Ch. 7 in *Asymptotics, Nonparametrics and Time Series*, S. Ghosh (Ed.). New York: Marcel Dekker, pp. 227–258.

Smart E. (1972). FKCOMB, A Fast General-Purpose Array Processor, *Seismic Array Analysis Center*, Teledyne Geotech, 20 December 1972.

Smart, E. (1976). Linear high-resolution frequency-wavenumber analysis, Ph,D. dissertation, Southern Methodist University.

Stoica, P. and A. Nehorai (1989). Music, maximum likelihood, and Cramer-Rao lower Bound, *IEEE Trans. Acoustics, Speech and Signal Processing* 37: 720–741.

Stroujkova, A. and D. Reiter (2006). Improving focal depth estimates: Studies of depth-phase detection at regional distances, in current Proceedings.

Tribuleac, I. M. and E. T. Herrin (1997). Calibration studies at TXAR, *Seis. Res. Lett.* 68: 353–365.